

Using Algorithms to Detect Gerrymandering and Improve Legislative Redistricting

Kosuke Imai

Harvard University

Roundtable on Civically Engaged Research, Democracy, and Civic Health

September 3, 2026

Joint work with many current and former Harvard undergraduate and graduate students
Chris Kenny (Princeton), Cory McCartan (Penn State), Tyler Simko (Michigan)



Motivation

- Today's world for **quantitative social science**:
 - ① increasing availability of granular data
 - ② rapid methodological and technological advancement including AI
- Social scientists can and should solve problems of the real world!
- **Redistricting** as a major policy problem
- How can we use data and algorithms to evaluate redistricting plans?
 - traditional methods: comparison across states and time periods
 - confounded by state-specific political geography and rules
- Use of **simulation algorithms**
 - ① obtain a representative sample of redistricting plans under constraints
 - ② compare the enacted plan with this baseline distribution
- Technological solution to detecting gerrymandering
- Tool for analyzing and improving redistricting

ALgorithm-Assisted Redistricting Methodology (ALARM)



ALARM Project

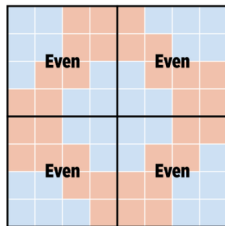
[Home](#) [About](#) [Applications](#)

**Developing methodology and tools
to analyze legislative redistricting.**

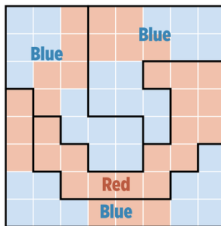
- What ALARM does:
 - ① develop efficient and flexible simulation algorithms
 - ② build open-source software packages for the entire workflow
 - ③ evaluate redistricting plans in the United States and elsewhere
- Goal: empower researchers, policy makers, data journalists, and citizen data scientists with powerful algorithmic tools

Redistricting Basics

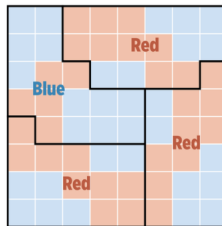
- Classic gerrymandering strategies: **packing** and **cracking**



Even distribution
2 red, 2 blue



Packing
1 red, 3 blue



Cracking
3 red, 1 blue

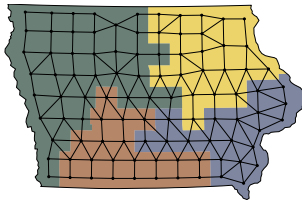
- What has changed:
 - availability of granular data
 - mapping software (e.g., Maptitude, Dave's Redistricting app)
- US Congressional redistricting
 - partisan gerrymandering: *Rucho v. Common Cause* (2019)
 - racial gerrymandering: *Louisiana v. Callais* (2026)

Why Use Simulation Algorithm for Redistricting Evaluation?

- Traditional redistricting evaluation: comparison across states and over time
- Confounded by differences in political geography and redistricting rules
- Simulation-based redistricting evaluation
 - ① generate many **alternative plans** under a set of redistricting criteria
 - ② compare them with a proposed plan to evaluate its properties
- Benefits of simulation approach
 - ① can control for **state-specific** political geography and redistricting rules
 - ② **transparency** and ability to isolate a relevant factor
 - ③ mathematical properties $\rightsquigarrow$ **representative sample** of alternative plans

Redistricting Simulation (Fifield et al. (2020). *Journal of Computational & Graphical Statistics*)

- Redistricting as a balanced graph partition problem



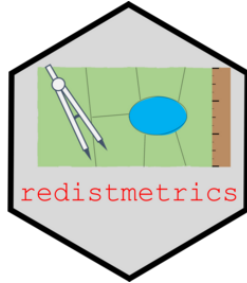
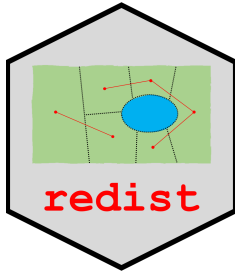
- Too many to enumerate
 - over 700 trillion ways to divide a 9×9 grid into 9 equal sized districts
 - obtain a random (and representative) sample from the population of valid redistricting plans
- Target distribution

$$\pi(\xi) \propto \underbrace{\tau(\xi)^\rho}_{\text{compactness}} \times \underbrace{\exp(-J(\xi))}_{\text{custom constraints}} \times \underbrace{1_{\xi \text{ connected}}}_{\text{contiguity}} \times \underbrace{1_{\text{dev}(\xi) \leq D}}_{\text{equal population}}$$

Sequential Monte Carlo Algorithm (McCartan and Imai. 2023. *Annals of Applied Statistics*)



Free and Easy-to-Use Open-source Software Development



50 State Redistricting Simulations Project (McCartan et al. 2022. *Nature Scientific Data*)

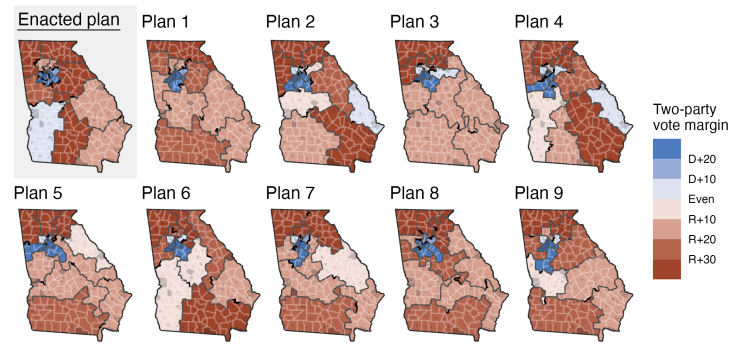


Comprehensive project to simulate alternative congressional redistricting plans for all fifty states.

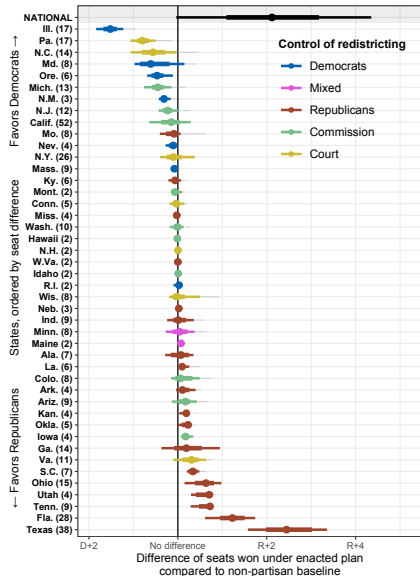
- tidied 2020 Census plus statewide election data from the VEST
- collect state-specific redistricting requirements
- construct algorithmic constraints based on these and traditional redistricting criteria
- simulated 5,000 nonpartisan plans
- code and data are available at Harvard Dataverse

Example: Georgia

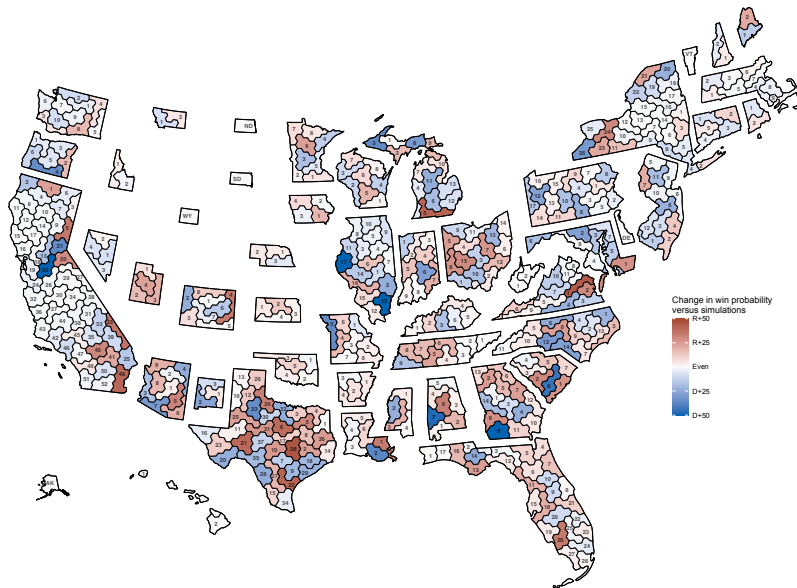
- House Legislative and Congressional Reapportionment Committee: districts must
 - 1 be contiguous
 - 2 have equal populations
 - 3 be geographically compact
 - 4 preserve county and municipality boundaries as much as possible
 - 5 avoid the unnecessary pairing of incumbents
- Account for everything except incumbency constraint: VRA is tricky



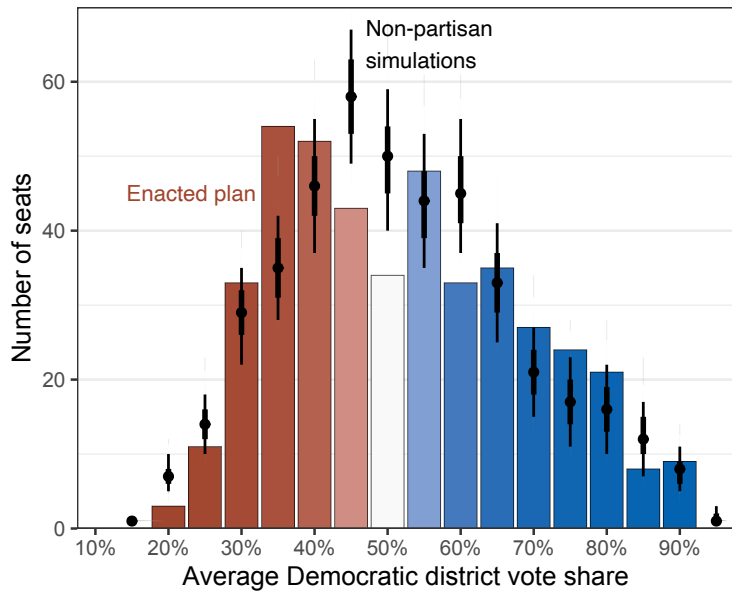
Widespread Partisan Gerrymandering Cancels Nationally (Kenny et al. 2023. *PNAS*)



Map of Partisan Gerrymandering



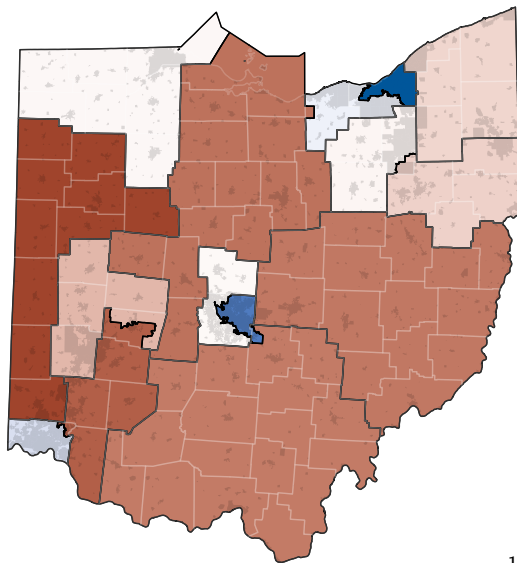
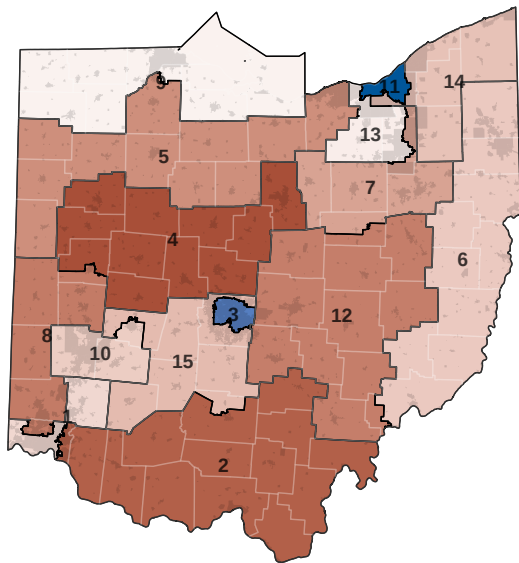
Partisan Gerrymandering Reduces Competitiveness



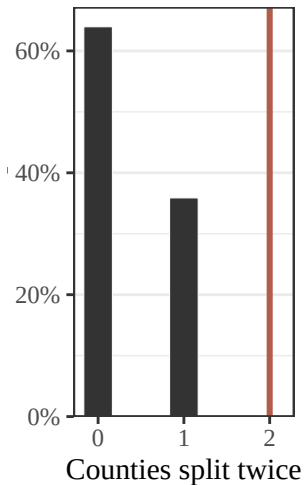
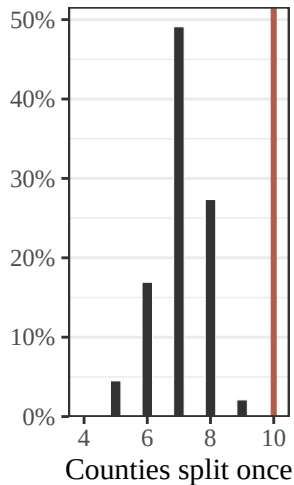
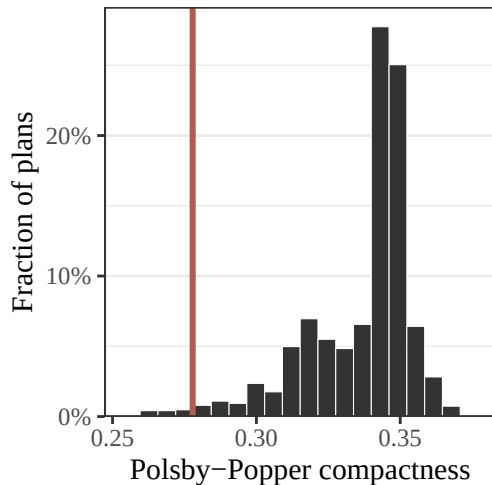
Application in the Court: Ohio Congressional Redistricting

- Previously 16 districts: 4 Democrats and 12 Republicans
- After 2020 Census, the number of seats is reduced to 15 districts
- 2018 Ohio voters passed the constitutional amendment
- I served as an expert witness for Relators: *League of Women Voters of Ohio et al. v. Ohio Redistricting Commission, et al.*
- Simulation analysis
 - 5,000 alternative plans
 - contiguous and compact districts
 - compliant with the Voting Rights Act (Cleveland)
 - Section 2(B)(5): out of Ohio's 88 counties,
 - at least 65 counties should not be split
 - no more than 18 counties can be split no more than once
 - no more than 5 counties can be split no more than twice

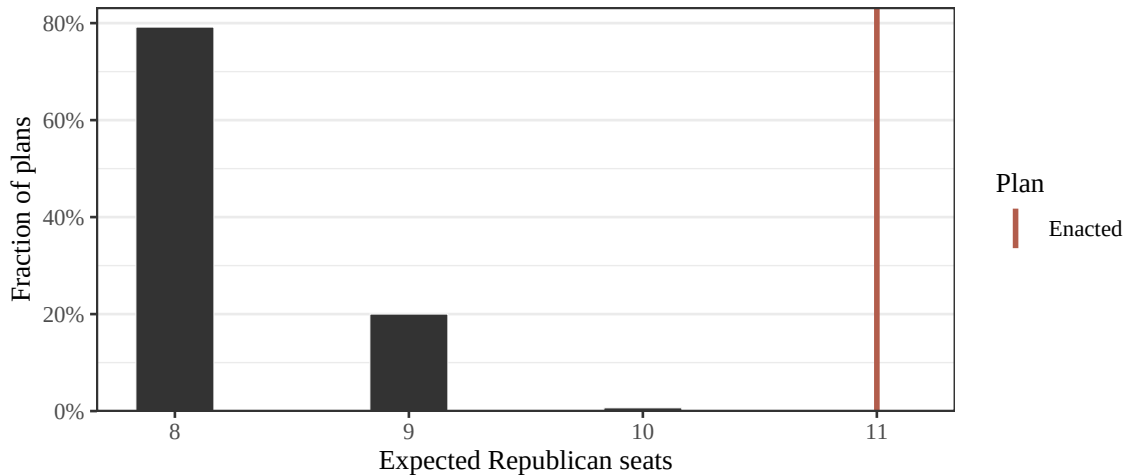
The Enacted and Example Simulated Plans



Compactness and Administrative Boundary Splits

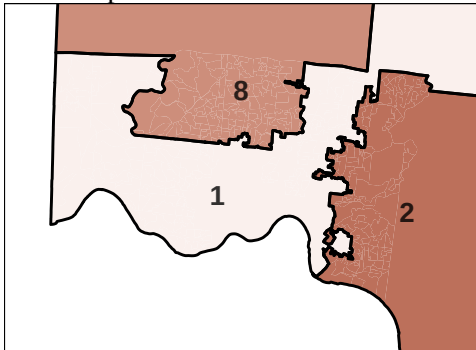


Expected Number of Republican Seats

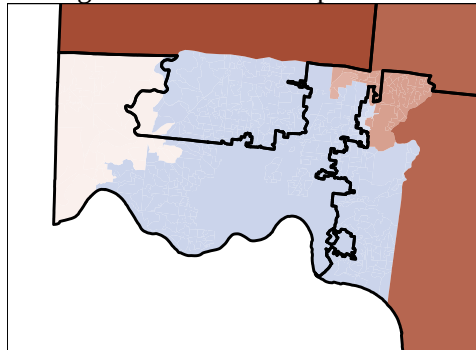


Cracking: Hamilton County (Cincinnati Area)

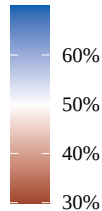
Enacted plan



Average across simulated plans

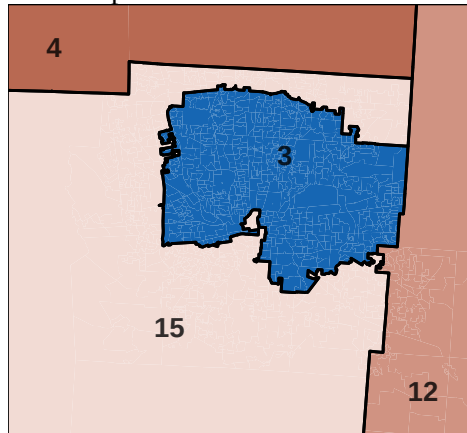


Two-party
vote share

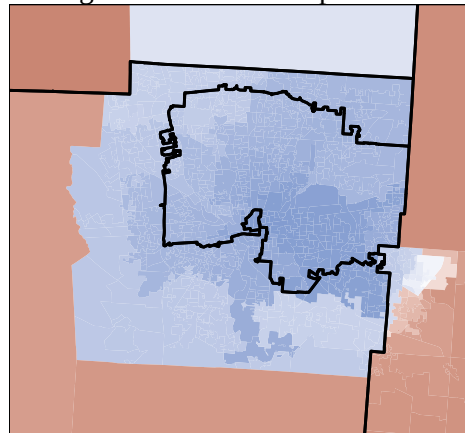


Packing: Franklin County (Columbus Area)

Enacted plan



Average across simulated plans



Two-party
vote share



60%
50%
40%
30%

Ohio Supreme Court Strikes Down the Enacted Map



The Court Opinion

Id. at Section 1(C)(3)(a). The above evidence, particularly Dr. Imai's conclusion that the enacted plan will result in, on average, 2.8 more Republican seats than are warranted, shows that the General Assembly's decision to shift what could have been—under a neutral application of Article XIX—Democratic-leaning areas into competitive districts, i.e., districts that give the Republican Party's candidates a better chance of winning than they would otherwise have had in a more compactly drawn district, resulted in a plan that unduly favors the Republican Party and unduly disfavors the Democratic Party.

Supreme Court: *Alexander v. South Carolina NAACP*.

- South Carolina racial gerrymandering case
- Served as an expert witness for the plaintiffs
- Used simulation to provide evidence that a disproportionately large number of Black voters are packed into District 6
- Federal district court ruled in favor of NAACP
- South Carolina directly appealed to the Supreme Court
- The Supreme Court reversed the federal district court's decision:

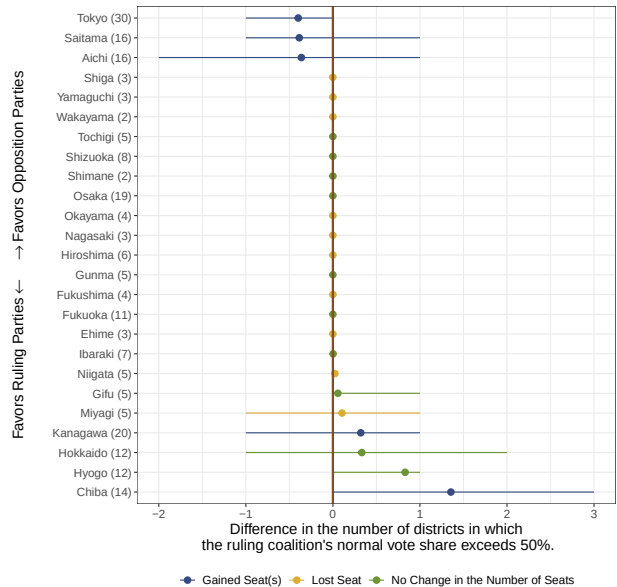
The Challengers' inference is flawed because Dr. Imai's models failed to consider partisanship. [...] Dr. Imai's algorithm produced maps without requiring that District 1 comply with the legislature's asserted aim of ensuring that District 1 remain a relatively safe Republican seat.

Beyond US Redistricting: Japan (Miyazaki et al. *Political Science Research and Methods* in-press)

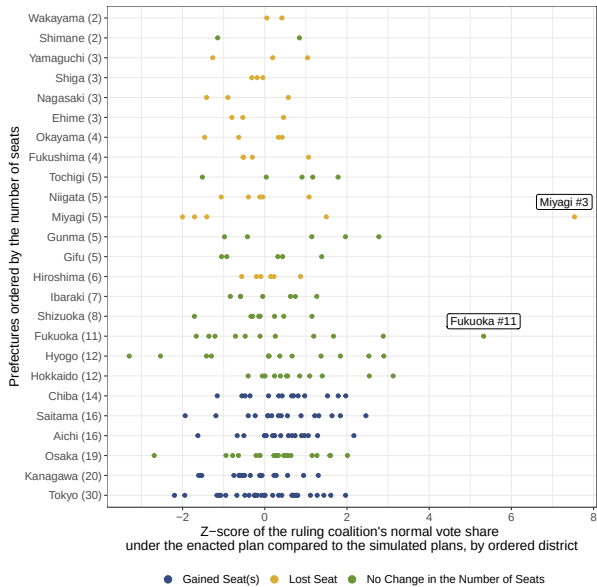
- Non-partisan commission $\leadsto$ no partisan bias?
- Potential sources of partisan bias
 - members are appointed by the prime minister and approved by the Diet
 - governors are invited to provide their opinions
- 2020 Japanese redistricting
 - redistricting in 25 prefectures out of 47
 - largest redistricting in the history



Little Partisan Bias at the Prefecture Level



Some but Relatively Little Partisan Bias at the District Level

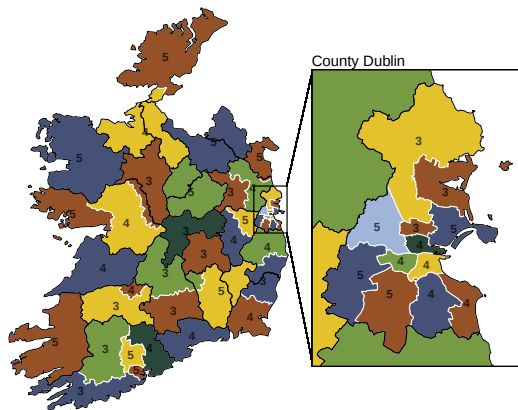


Ireland Redistricting (O'Sullivan et al. *Working paper*)

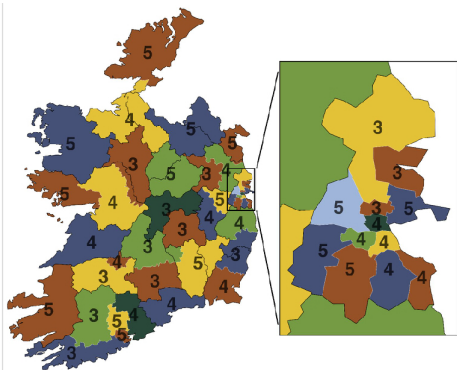
- Ongoing collaboration with the Irish Electoral Commission



- Dáil Éireann (Lower House) redistricting
 - multimember district: 3, 4, or 5 seats per district
 - 171–181 seats in total
- We generalized our algorithm to handle multimember districts



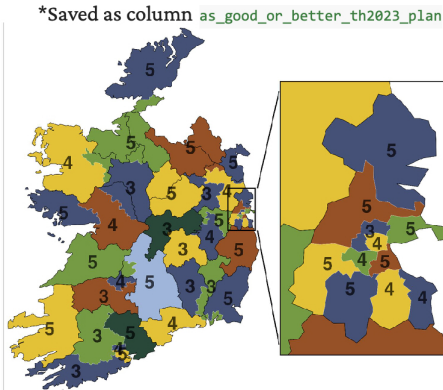
It Is Possible to Improve the 2023 Plan



- 43 Districts
- 8 Splits
- Maximum Deviation $\approx 8.12\%$ (per seat)
- Similarity $\approx 85.50\%$

Constituency Sizes

3	4	5
13	15	15

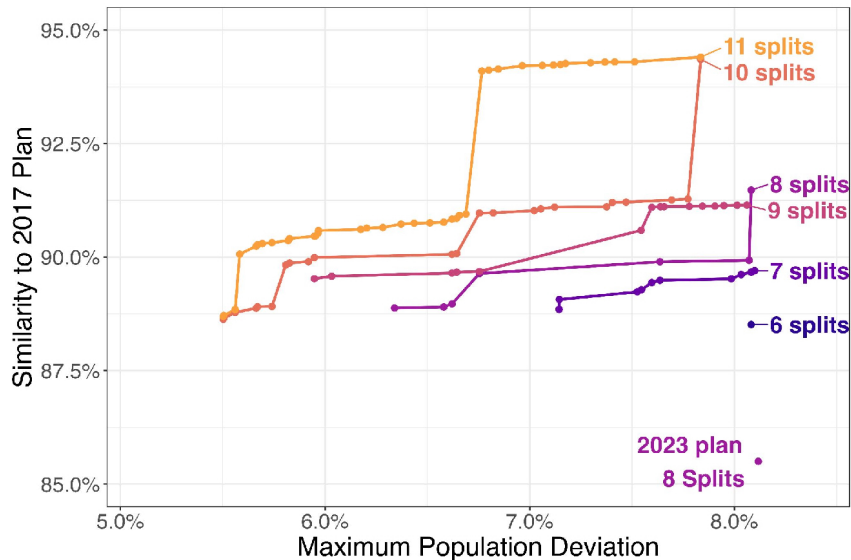


- 41 Districts
- 8 Splits
- Maximum Deviation $\approx 6.34\%$ (per seat)
- Similarity $\approx 88.88\%$

Constituency Sizes

3	4	5
10	11	20

Understanding Tradeoffs between Different Criteria



Concluding Remarks

- Redistricting matters
 - fair representation and policy outcomes
 - competitiveness of districts and responsiveness
 - geographic polarization
- How should we stop gerrymandering?
 - independent commissions
 - electoral reforms
- Role of academic experts
 - testimony to courts and legislatures
 - collaboration with non-partisan groups and commissions
- Ongoing work
 - impact of commission and other reforms (*APSR*, in-press)
 - analysis of the *Louisiana v. Callais* decision (working paper)
 - analysis of Congressional redistricting in 1990 and 2000
 - analysis of State House and Senate redistricting